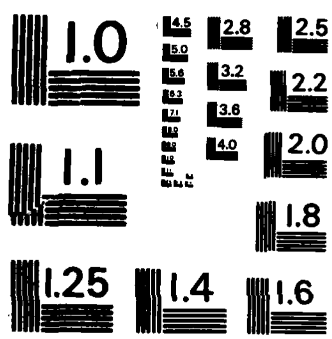
AD-A166 151 TAIL BEHAVIOR FOR THE SUPREMA OF GAUSSIAN PROCESSES 1/1
WITH A VIEW TOWARDS E. (U) NORTH CAROLINA UNIV AT
CHAPEL HILL CENTER FOR STOCHASTIC PROC
UNCLASSIFIED R J ADLER ET AL NOV 85 TR-127 F/G 12/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR-TR- 86-0015

(2)

CENTER FOR STOCHASTIC PROCESSES

AD-A166 151

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina

TAIL BEHAVIOUR FOR THE SUPREMA OF GAUSSIAN PROCESSES WITH A VIEW TOWARDS EMPIRICAL PROCESSES

by

Robert J. Adler

and

Gennady Samorodnitsky

Technical Report No. 127

November 1985

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY 2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; Unlimited distribution unlimited.
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Report No. 127	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 86-0015
6a. NAME OF PERFORMING ORGANIZATION Center for Stochastic Processes 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research
6c. ADDRESS (City, State and ZIP Code) Statistics Dept., Univ. of North Carolina Phillips Hall 039-A Chapel Hill, NC 27514	7b. ADDRESS (City, State and ZIP Code) Bolling Air Force Base Washington, DC 20332
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR 8b. OFFICE SYMBOL (If applicable) N/M	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F49620-85-C-0144
8c. ADDRESS (City, State and ZIP Code) Bolling Air Froce Base Washington, DC 20332	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO.: 61102F PROJECT NO.: 2304 TASK NO.: A5 WORK UNIT NO.:

11. TITLE (Include Security Classification)
"Tail Behaviour for the suprema of Gaussian processes with a view towards empirical processes"

12. PERSONAL AUTHOR(S)
R.J. Adler and G. Samorodnitsky

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
technical	FROM 9/85 TO 8/86	Nov. 85	50

16. SUPPLEMENTARY NOTATION
Gaussian processes, isonormal process, supremum, metric entropy, Brownian sheet, empirical processes.

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Initially we consider "the" standard isonormal linear process L on a Hilbert space H, and applying metric entropy methods obtain bounds for the probability that $\sup_C Lx > \lambda$, $C \subset H$ and λ large. Under the assumption that the entropy function of C grows polynomially, we find bounds of the form $c\lambda^{\alpha}e^{-1/2\lambda^2/\sigma^2}$, where σ^2 is the maximal variance of L. We use a notion of entropy finer than that usually employed, and specifically suited to the non-stationary situation. As a result we obtain, in the non-stationary setting, more precise bounds than any in the literature.

We then treat a number of examples in which the power α is identified. These include the distribution of the maximum of certain "locally stationary" process on $\mathbb{R}^1$, as well as those of the rectangle indexed, pinned Brownian sheet on $\mathbb{R}^k$, for which $\alpha=2(2k-1)$, and the half-plane indexed pinned sheet on $\mathbb{R}^2$ for which $\alpha = 2$.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL ~~Peggy Ravitch~~ Jacobs	22b. TELEPHONE NUMBER (Include Area Code) (202) 767-4940 22c. OFFICE SYMBOL NM

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

TAIL BEHAVIOUR FOR THE SUPREMA OF GAUSSIAN PROCESSES WITH A VIEW TOWARDS EMPIRICAL PROCESSES

Robert J. Adler[1]
Faculty of Industrial Engineering & Management
Technion-Israel Institute of Technology
and
Center for Stochastic Processes
Statistics Department
University of North Carolina at Chapel Hill

and

Gennady Samorodnitsky[2]
Faculty of Industrial Engineering & Management
Technion-Israel Institute of Technology

[1]Research supported in part by AFOSR Constract Nos. 84-0104, 85-0384 and F49620 85 C 0114 while visiting the Center for Stochastic Processes, Chapel Hill, North Carolina.

[2]Research supported in part by the Wolf foundation.

A-

SUMMARY

Initially the document ~~we~~ considers "the" standard isonormal linear process L on a Hilbert space H, and applying metric entropy methods obtain bounds for a certain ~~the~~ probability, that $\sup_C Lx > \lambda$, $C \subset H$ and λ large. Under the assumption that the entropy function of C grows polynomially, we find bounds of the form $c\lambda^{\alpha}e^{-\frac{1}{2}\lambda^2/\sigma^2}$, where delta squared σ^2 is the maximal variance of L. We use a notion of entropy finer than that usually employed, and specifically suited to the non-stationary situation. As a result we obtian, in the non-stationary setting, more precise bounds than any in the literature.

We then treat a number of examples in which the power alpha α is identified. These include the distribution of the maximum of certain "locally stationary" process on $\mathbb{R}^1$, as well as those of the rectangle indexed, pinned Brownian sheet on $\mathbb{R}^k$, for which $\alpha=2(2k-1)$, and the half-plane indexed pinned sheet on $\mathbb{R}^2$ for which $\alpha = 2$.

AMS 1980 subject classifications
Primary 60G15, 60G57. Secondary 60F10, 62G30.

Key words and phrases
Gaussian processes, isonormal process, supremum, metric entropy, Brownian sheet, empirical processes.

Running head Suprema of Gaussian Processes

1. INTRODUCTION

We start with some motivation from the theory of empirical processes, letting $X_1,\dots,X_n$ be i.i.d. observations from some k-dimensional distribution, and assuming we want to test the hypothesis that the parent distribution is given by a measure ν: $\nu(A) = P\{X_i \in A\}$ on the unit cube. A natural test procedure is to form the empirical measure ν_n: $\nu_n(A) = \frac{1}{n}\sum_{i=1}^{n} I_A(X_i)$ (I_A is the indicator function of A) and compare ν_n to ν via a Kolmogorov-Smirnov type statistic of the form

$$\sup_A \{\sqrt{n}\, |\nu_n(A) - \nu(A)|\} \tag{1.1}$$

for some family $\mathcal{A}$ of Borel subsets of $[0,1]^k$. It is known (Dudley 1978, 1984) that $\sqrt{n}(\nu_n-\nu)$ converges weakly to a Gaussian process on $\mathcal{A}$, under conditions related to the size of $\mathcal{A}$. Consequently, the study of (1.1) reduces, in the limit, to the study of the supremum of a particular Gaussian process over a class of sets.

Unlike the case for their Markov counterparts, however, it is well known that for Gaussian processes it borders on the impossible to obtain the exact distribution of their (global) maxima. For stationary Gaussian processes on the line, for example, there are only six covariance functions for which the precise distribution of the maxima of the corresponding processes are known (c.f. Slepian (1961), Slepian and Shepp (1976), Cressie and Davis (1981), Darling (1983)). For random fields on $\mathbb{R}^k$ the situation is even worse, for there exists no non-trivial Gaussian field, either stationary or not, for which the precise distribution of the maxima is known. In certain specific cases, however, upper and lower bounds to this distribution are known.

Goodman (1976), for example, calculated good bounds for the cases of the pinned and regular Brownian sheets in $\mathbb{R}^2$. (See Section 4 for definitions).

These have been improved and extended to higher dimensions in Cabaña and Wschebor (1982), Cabaña (1984) and Adler and Brown (1986). All but the last reference deal only with sheets arising from the case ν = Lebesgue measure in (1.1). The only other Gaussian field for which some (not wholly satisfactory) bounds are known is a two-parameter generalisation of Slepian's triangular covariance function (Cabaña and Wschebor (1981), Adler (1984)).

Needless to say, in more general situations, such as those arising from (1.1) when the parameter space may be a class of sets, virtually nothing is known on the exact distribution of the supremum.

Partly, or perhaps primarily, because of this dirth of results a large amount of effort has been expended in studying the asymptotic properties of Gaussian maxima. The most central, and most well known result in this direction is due to four authors, Fernique (1970, 1975), Landau and Shepp (1971) and Marcus and Shepp (1971), who proved various versions of the result that for any zero mean sample path continuous Gaussian process $X(t)$, $t \in S$, and S a metric space,

$$\lim_{\lambda\to\infty} \frac{\ln P\{\sup(X(t),\ t \in S) > \lambda\}}{\lambda^2} = -1/(2\sigma^2) \tag{1.2}$$

where

$$\sigma^2 = \sup_{t\in S} E\{X^2(t)\}.$$

An immediate consequence of (1.2) is that for all $\lambda_0 > 0$, and any $\varepsilon > 0$, there exists a constant $K=K(\varepsilon,\lambda_0)$ such that if $\lambda > \lambda_0$ then

$$P\{\sup_{t\in S} X(t) > \lambda\} \le K\, e^{\varepsilon\lambda^2} e^{-\frac{1}{2}\lambda/\sigma^2} \tag{1.3}$$

(An even sharper result than this is due to Borell (1975). See comment 3 of Section 6.)

Our aim in this paper will be to perform a simple epsilonectomy - i.e. to remove the factor $\exp(\varepsilon\lambda^2)$ from (1.3). In general this cannot be done without paying some price, and in the cases we shall consider the price will be to replace this exponential factor by a smaller power factor of the form λ^α, $\alpha \geq -1$, so as to obtain bounds of the form

$$(1.4) \qquad P\{\sup_{t\in S} X(t) > \lambda\} \leq k\lambda^\alpha e^{-\frac{1}{2}\lambda^2/\sigma^2},$$

for large enough λ.

Results like (1.4) are not new. They were obtained originally by Pickands (1969a,b) for the class of zero mean, stationary Gaussian processes on [0,1] whose covariance function $R(t) = E\{X(s)X(s+t)\}$ satisfies

$$(1.5) \qquad R(t) = 1 - c|t|^\alpha + o(|t|^\alpha) \qquad \text{as } |t| \to 0,$$

where $\alpha \in (0,2]$ and $c > 0$ are constants. Pickands showed that for each fixed $h > 0$ for which $\sup_{\varepsilon\leq t\leq h} R(t) = \delta_\varepsilon < 1$ for all $\varepsilon > 0$

$$(1.6) \qquad \lim_{\lambda\to\infty} \frac{1}{\lambda^{2/\alpha} p(\lambda)/\lambda} P\{\sup_{0\leq t\leq h} X_t > \lambda\} = hC^{1/\alpha} H_\alpha,$$

where $H_\alpha > 0$ is a finite constant depending only on α and p is a standard normal density function. (Except for the cases $\alpha = 1$, $\alpha = 2$, the value of H_α is not known.) This result has been extended to certain stationary random fields by Belyaev and Piterbarg (1972) and, more recently, to certain non-homogeneous processes on $\mathbb{R}^1$ by Piterbarg and Prisjažnjuk (1979). A proof of (1.6), along with historical details, can be found in Leadbetter, Lindgren and Rootzen (1983).

More recently Weber (1978, 1980) has obtained a set of results which, while they do not identify constants as in (1.6), provide bounds to the distributions of Gaussian suprema for the widest possible class of Gaussian processes, including the set-indexed processes described above. However, as we shall show later, his bounds, when they are of the form of (1.4), do not always yield the smallest possible value of α. We shall have more specific comments to make about Weber's results later.

Before saying any more, it is probably worthwhile at this point to explain to the sceptic what we gain from an epsilonectomy at (1.3) beyond the surgeon's natural pleasure of neatly removing an unnecessary appendage or, indeed, from sharpening the power in Weber's results. The first application is purely theoretical. Consider a function valued Gaussian process, i.e. a process $Y_{(\cdot)}$, whose value at a given time is a Gaussian random process. Such processes arise naturally in a number of ways, often by "relabelling", for example, a two-parameter process $X(s,t)$ to obtain a function valued Y_t under the correspondence $Y_t(s) = X(s,t)$. Such processes include the Kiefer process (Kiefer (1972)) of empirical process theory. Iterated logarithm type results for the growth of $\sup_s Y_t(s)$ with t have been studied in depth (see, for example, Goodman, Kuelbs and Zinn (1981)) and, to a heavy extent, are based on the inequality (1.3). Finer results, such as upper-lower class theorems for $\sup_s Y_t(s)$, are much harder to obtain (Kuelbs, (1975) is one exception we are aware of) as (1.3) does not provide fine enough information. A result of the form (1.4) does, however, fulfill this need, and is applied to this purpose to obtain upper-lower class theorems for empirical processes in Adler and Brown (1986). Establishing (1.4) in general, therefore, opens up the possibility of a general upper-lower class theory for function valued processes.

For the second application we return to our opening paragraph and the Kolmogorov-Smirnov type statistic (1.1). Although our results will bound the (asymptotic in n) tail distribution of (1.1), they will not really do so sharply enough to enable, say, the generation of critical levels for statistical tests. This problem seems to be hard enough that for the foreseeable future this will be done by simulation techniques. What a bound like (1.4) tells the simulator, however, is that the critical levels depend on three parameters, k,α, and σ^2. As will be shown in Section 4, α and σ^2 can be obtained from our general theory, so that only one parameter remains to be estimated, making the simulation task much simpler.

The paper is organized as follows. In order to treat the most general processes possible, we shall work initially with the isonormal Gaussian process on Hilbert space. This, together with requisite entropy notions, will be described in the following section, where we shall also develop a version of Fernique's (1975) inequality, that will be the basis of all that follows. In Section 3 we shall present a number of theorems that show that by putting more and more structure on the parameter Hilbert space (via entropy conditions) finer and finer bounds on the distribution of the maximum can be obtained. Proofs are deferred to Section 5. Section 4 contains a number of examples, in which we apply the results on the isonormal process to specific problems. For example, we obtain sharp (in the sense of best possible power α) bounds for the maximum of a rectangle indexed Brownian sheet. In Section 6 we conclude with some comments.

Acknowledgements. Some of the results presented here, when restricted to the class of homogeneous Gaussian fields on $\mathbb{R}^k$, have a significant overlap with the "extended Fernique inequality" in Berman (1985a). We had already obtained these results independently before hearing, from Professor Berman, of this work. However, when he very kindly sent us a preliminary (still untyped) version of his results we took advantage of the opportunity to combine what was best in both proofs, and so the statements and proofs of Theorems 3.2 and 3.3, when restricted to simple random fields, have much in common with his results. As our examples show, however, even for simple fields, our later theorems go beyond his in identifying the optimal power.

We are also grateful to Larry Brown, who did most of the hard work in Adler and Brown (1986). It was his insight on the problems tackled there that set us off on the current work.

Both a referee, and Professor Weber himself, drew our attention to the results of Weber (1978, 1980). We are grateful to Professor Weber for correspondence helping to clarify the relationships between his work and an earlier version of this paper.

2. THE ISONORMAL PROCESS AND A FERNIQUE INEQUALITY

The central idea is to study one, canonical, Gaussian process, and then relate any particular process to this one. It is defined as follows. Call a sequence $\{X_n\}$ of random variables orthogaussian iff they are independent with $L(X_j) \equiv N(0,1)$. Let H be a real, infinite-dimensional Hilbert space. A linear map L from H into real Gaussian variables with $EL(x) = 0$ and $EL(x)L(y) = (x,y)$ for all $x,y \epsilon H$ is called the isonormal Gaussian process on H. (c.f. Segal (1954), Dudley (1967, 1973). For example, if $\{x_n\}$ is an orthonormal basis for H so that for $x \epsilon H$, $x = \Sigma a_n x_n$, we can let $L(x) = \Sigma a_n Y_n$, where the Y_n are orthogaussian.

Since Gaussian distributions are uniquely determined by their means and covariances, the isonormal process L can be regarded as the only real Gaussian process. For, if $\{x_t, t\epsilon T\}$ is any real Gaussian process with mean $Ex_t = m_t$, then $L(x_t - m_t) + m_t$ is another version of the process, where we take $L^2(\Omega,P)$ for H. On H,L "remembers" the covariance structure of x_t, and, by its linearity, also keeps track of all joint distributions. Thus, we can in general neglect the specific joint distributions of x_t on (Ω,P) and work only with the abstract geometric structure of the function $t \to x_t - m_t \epsilon H$. To see precisely how this works in practice, see the examples in Section 4.

In order to study the structure of H, we shall require the notion of metric entropy. Let C be a subset of a metric space (S,d). Given $\varepsilon>0$, let $N(C,\varepsilon) \equiv N_C(\varepsilon)$ be the minimal number of points $x_1,\ldots,x_n$ from C such that for all $y\epsilon C$ $\min_i d(x_i,y) \leq \varepsilon$. We assume N finite for all $\varepsilon > 0$. Consequently, there exist sets $A_1,\ldots,A_{N_C(\varepsilon)}$ covering C such that for all n $d(x,y) \leq 2\varepsilon$ for all $x,y\epsilon A_n$. Set $H_C(\varepsilon) = \log N_C(\varepsilon)$.

Then $H_C(\varepsilon)$ is the metric entropy of C. Metric entropy is well known to play an important role in continuity problems for Gaussian processes. For example, L, restricted to $C \subset H$, is sample continuous if $\int_0^1 H_C^{1/2}(x)\,dx < \infty$. Metric entropy can also be used to study suprema problems. For example, Weber (1980) has shown that if $||x|| = 1$ for all $x \in C$, and certain other side conditions hold, then

$$\text{(2.1)} \qquad P\{\sup_{x \in C} |Lx| > \lambda + \Pi_\lambda\} \leq \text{const. } N(C, v(\lambda))\Psi(\lambda),$$

where

$$\Psi(\lambda) = P\{|Lx| > \lambda\} = \sqrt{2/\pi} \int_\lambda^\infty e^{-\frac{1}{2}u^2}\,du,$$

$$\Pi_\lambda = \rho(\rho-1) \int_0^{\rho v(\lambda)} [H(C,\varepsilon) - \log \varepsilon]^{-\frac{1}{2}}\,d\varepsilon,$$

$$v(\lambda) = \inf \{0 < \varepsilon < \varepsilon_0 \colon h(\varepsilon) \leq \lambda\},$$

$$h(\varepsilon) = \varepsilon^{-1} [H(C,\varepsilon) - \log \varepsilon]^{\frac{1}{2}},$$

$$\varepsilon_0 = \inf \{0 < \varepsilon < 1 \colon N(C,\varepsilon) < 2\}$$

and $\rho \in (0,1)$ is arbitrary. Assuming Π_λ is small enough for large λ (as is usually the case), that v is at most polynomial, and that the entropy is polynomial, we see that (2.1) is a result of the form of (1.4), which is what we are seeking.

There are, however, two difficulties with Weber's result, insofar as general best upper bounds are concerned, and, in particular in relation to the examples from the theory of empirical processes that motivated us. The first is the assumption that $||x|| = 1$ for all x. It is possible to get around this in the general case by noting

$$(2.2) \qquad P\{\sup_{x\epsilon C} Lx > \lambda\} < P\{\sup_{y\epsilon C'} Ly > \frac{\lambda}{\sigma}\}$$

where $\sigma = \sup ||x||$, and $C' = \{y: y = x/||x||, x \in C\}$. It is not hard to see that the entropy function for C' follows the same general behaviour of that for C, and since $||y|| = 1$ for $y \in C'$ Weber's result then gives a bound for (2.2). However, it is easy to check via examples such as Example 4.1 that this procedure does not give the sharpest bounds possible.

The second difficulty to somewhat more fundamental, and essentially insurmountable, even if Weber's results did not assume $||x|| = 1$. It lies in the fact that a methodology based purely on metric entropy can never always give the best bounds. To see this, one example will suffice. In Section 4 we show how to calculate supremum distributions for general processes by assigning to each process a particular Hilbert space, and then studying L on that space. It is easy to see that the Wiener process, $W(t)$, $t \in [0,2]$ and the stationary Slepian process $S_t := W_{t+1} - W_t$, $t \in [0,1]$ generate identical (up to a constant) entropy functions since

$$E\{|W_t - W_s|^2\} = |t - s| = \tfrac{1}{2}E\{|S_t - S_s|^2\}, \qquad 0 < s,t < 1.$$

Thus any bound for the suprema distributions of W and S on [0,1] coming from metric entropy considerations involving only H must be the same. But it is well known that whereas $P\{\sup_{[0,1]} W_t > \lambda\} = O(\lambda^{-1}e^{-\frac{1}{2}\lambda^2})$, we have $P\{\sup_{[0,1]} S_t > \lambda\} = O(e^{-\frac{1}{2}\lambda^2})$.

In general, then, the problem is that different processes may have essentially the same metric entropy, but quite different suprema distributions.

In order to solve this problem we shall require finer partitions on C than those obtainable just from entropy considerations. To this end, for given $\delta \geq 0$ set

$$(2.1) \qquad C_\delta^+ = \{x\epsilon C: ||x|| > \delta\}, \quad C_\delta^- = \{x\epsilon C: ||x|| \leq \delta\},$$

where $C \subset H$ and $||.||$ is the H-induced norm. Now define

$$(2.2) \qquad N_C^+(\delta,\varepsilon) := N(C_\delta^+,\varepsilon), \; N_C^-(\delta,\varepsilon) := N(C_\delta^-,\varepsilon).$$

Since $C = C_\delta^+ \cup C_\delta^-$, it is obvious that $N_C(\varepsilon) \leq N_C^+(\delta,\varepsilon) + N_C^-(\delta,\varepsilon)$ for all δ and ε. We shall need one more entropy function,

$$(2.3) \qquad N_C(\delta_1,\delta_2,\varepsilon) := N(C_{\delta_1}^+ \cap C_{\delta_2}^-,\varepsilon), \quad 0 \leq \delta_1 \leq \delta_2, \; \varepsilon > 0.$$

The motivation behind this last entropy function should be clear. The idea is to first break up C into regions over which $L(x)$ has a variance $(=||x||^2)$ within certain bounds, and then to measure the "size" of each of these regions via entropy considerations. This will provide the finer information we shall need (particularly for non-homogeneous processes for which $||x||$ is not constant over H) to obtain sharp bounds for the distribution of $\sup L(x)$.

We can now commence setting up the basic (Fernique type) inequality from which all our other results will ultimately follow. To this end, set

$$\sigma = \sup_{x\epsilon C} ||x||.$$

Let δ_i be a sequence satisfying $0 = \delta_0 < \delta_1 < \ldots < \delta_m = \sigma$, with m possibly infinite. For each $i=1,\ldots,m$ let ε_{ij}, $j=1,2,\ldots$, be an infinite monotone sequence such that $\lim_{j\to\infty} \varepsilon_{ij} = 0$. We shall use these two sequences to partition C as the union of $C(\delta_{i-1},\delta_i)$, where

$$C(\nu,\eta) := C_\nu^+ \cap C_\eta^- = \{x \epsilon C: \nu < \|x\| \leq \eta\},\ 0 \leq \nu < \eta \leq \sigma.$$

Note that for every j there is a finite collection of points of $C(\delta_{i-1},\delta_i)$, which we shall denote by C_{ij}, satisfying

(2.4) $\quad \#C_{ij} = N_C(\delta_{i-1},\delta_i,\varepsilon_{ij})$,

(2.5) $\quad$ for all $y \epsilon C(\delta_{i-1},\delta_i)$ there exists an $x \epsilon C_{ij}$ such that $\|x-y\| < \varepsilon_{ij}$.

(Here #A is the cardinality of A.)

We shall need one more double sequence λ_{ij}, $i=1,\ldots,m$, $j=0,1,2,\ldots$, of positive numbers. Clearly

$$P\{\sup_{x\epsilon C_{i1}} |Lx| > \lambda_{i0}\delta_i\} \leq N_C(\delta_{i-1},\delta_i,\varepsilon_{i1})\psi(\lambda_{i0}), \tag{2.6}$$

where

$$\psi(u) = \sqrt{2/\pi}\int_u^\infty e^{-\frac{1}{2}x^2}dx\ . \tag{2.7}$$

Furthermore, for each $x \epsilon C(\delta_{i-1},\delta_i)$ there is a point $x_{ij}(x) \epsilon C_{ij}$ such that $\|x-x_{ij}\| < \varepsilon_{ij}$. Consequently

$$P\{\sup_{x\in C_{i,j+1}} |Lx - Lx_{ij}(x)| > \lambda_{ij}\varepsilon_{ij}\} \leq N_C(\delta_{i-1},\delta_i,\varepsilon_{i,j+1})\psi(\lambda_{ij}),$$

from which follows that

$$(2.8)\qquad P\{\sup_{x\in C_{i,j+1}} |Lx| > \lambda_{i0}\delta_i + \sum_{k=1}^{j} \lambda_{ik}\varepsilon_{ik}\}$$

$$\leq \sum_{k=0}^{j} N_C(\delta_{i-1},\delta_i,\varepsilon_{i,k+1})\psi(\lambda_{ik}) .$$

Now note that, as $j \to \infty$, C_{ij} becomes dense in $C(\delta_{i-1},\delta_i)$. Consequently, choosing a separable version of L we obtain from (2.8) that

$$P\{\sup_{C(\delta_{i-1},\delta_i)} |Lx| > \lambda_{i0}\delta_i + \sum_{j=1}^{\infty} \lambda_{ij}\varepsilon_{ij}\} \leq \sum_{j=0}^{\infty} N_C(\delta_{i-1},\delta_i,\varepsilon_{i,j+1})\psi(\lambda_{ij}).$$

It is now trivial to check the truth of the following inequality, which forms the basis of the remainder of the paper.

<u>Basic Inequality</u> <u>For sequences</u> δ_i, λ_{ij} <u>and</u> ε_{ij} <u>satisfying</u> $0 = \delta_0 < \delta_1 < \dots < \delta_m = \sigma$ (m <u>possibly infinite</u>) <u>and</u> $\varepsilon_{ij} \searrow 0$ <u>as</u> $j \to \infty$ <u>for all</u> i, <u>separable versions of</u> L <u>satisfy</u>

$$(2.9)\qquad P\{\sup_{x\in C} |Lx| > \sum_{i=1}^{m} \lambda_{i0}\delta_i + \sum_{i=1}^{m}\sum_{j=1}^{\infty} \lambda_{ij}\varepsilon_{ij}\}$$

$$\leq \sum_{i=1}^{m}\sum_{j=0}^{\infty} N_C(\delta_{i-1},\delta_i,\varepsilon_{i,j+1})\psi(\lambda_{ij}) .$$

Note that this basic estimate is extremely general, and not particularly informative. Our task now will be to propose meaningful, checkable conditions on $N_C(\nu,\eta,\varepsilon)$, and, by judicious choices of the various sequences in the basic inequality, reduce the various sums in (2.9) to simple, useful, forms.

3. MAIN RESULTS

There are basically two types of possible growth rates for entropy functions that yield interesting results on sup Lx, polynomial growth of the form $N_C(\varepsilon) \sim a\varepsilon^{-\kappa}$, or exponential growth of the form $N_C(\varepsilon) \sim a \exp(\varepsilon^{-\kappa})$. Faster than exponential growth rates yield discontinuous, unbounded processes for which no non-trivial bound on the distribution of sup L can exist, and slower than power rates are generally just not interesting. In this paper we shall study only polynomial entropies, and shall show how to relate the κ above to the α of (1.4). For some remarks on exponential entropies, see Section 6.

Polynomial entropies, while initially seemingly restrictive, cover a wide range of examples, including random fields indexed by finite dimensional Euclidean space and processes indexed by spaces of sets, such as polygons, that are describable by a finite number of parameters. Processes indexed by Vapnik-Cervonenkis classes of sets or functions (c.f. Section 6) are also described by polynomial entropies. (c.f., for example, Dudley (1973, 78, 84).)

For the first result, we shall assume only minimal information on C, which also turns out to be all that is required if L is stationary on C (implied by $\|x\|$ = const. for all $x \in C$ and $(x,y)=f(x-y)$ for all $x,y \in C$ and some positive definite f). To be more precise, we assume there exist positive constants a and κ such that

$$(3.1) \qquad N_C(\varepsilon) \equiv N_C(0,\sigma,\varepsilon) \leq a\varepsilon^{-\kappa}$$

for small enough ε. Then it is easy to show via the basic inequality (2.9) (c.f. Section 5) that for large enough $p \geq 2$ and all $\lambda > (1+4\kappa \ell np)^{\frac{1}{2}}$

$$\text{(3.2)} \qquad P\{\sup_{x\in C} |Lx| > \lambda(\sigma+2p^{-2})\} \le \frac{5}{2}\, ap^{2\kappa}\int_\lambda^\infty e^{-\frac{1}{2}u^2}du\ .$$

To the reader acquainted with Fernique (1975) this inequality should appear familiar, for he has a similar inequality for processes on Euclidean space. It is in fact a simple matter to derive Fernique's inequality from (3.2).

Via (3.2) it is not hard to prove the following result, closely related to Théorème 2.1 of Weber (1980) in the case $||x|| = \sigma = 1$ for all x.

Theorem 3.1 Suppose $N_C(\varepsilon) \le a\varepsilon^{-\kappa}$ for all $\varepsilon\in(0,\varepsilon_0]$. Define the following constants.

$$b = b(\kappa,\varepsilon_0) = \begin{cases} \max(\varepsilon_0^{-\frac{1}{2}},2,2\kappa+1) & 0<\kappa<4, \\ \max(\varepsilon_0^{-\frac{1}{2}},2,1+2\sqrt{\kappa}\,\ln\kappa) & \kappa\ge 4, \end{cases}$$

$$M_1 = \frac{5}{2}a(\sigma+\tfrac{1}{2}),\ M_2 = \frac{5}{2}a(\sigma+\tfrac{1}{2})\exp\{(2\sigma+\tfrac{1}{2})/\sigma^4\}\ .$$

Then, for all $\lambda \ge 2b(\sigma+\frac{1}{2})^2$,

$$\text{(3.3)} \qquad P\{\sup_{x\in C}|Lx| > \lambda\} \le M_1\lambda^{2\kappa-1}e^{-\lambda^2/2\sigma^2}.\ \exp\{2(\sigma+\lambda^{-2})/\sigma^4\}$$

$$\le M_2\lambda^{2\kappa-1}e^{-\lambda^2/2\sigma^2}.$$

Two things should be noted about this result. The first is that since the assumptions assume nothing about the variation of $||x||$ on C, (3.3) is unlikely to lead to sharp bounds for non-homogeneous processes. In fact, it doesn't. Secondly, the constants in (3.3), while a little unwieldy, are identifiable. As we assume finer structure on C, while we shall get smaller powers for the power of λ in (3.3), we shall lose track of the constants. (In principle, we could always keep track of the constants, but one reaches a point where they become so complicated that it no longer seems worthwhile to expend the not inconsiderable effort required to do so.)

Our first step away from homogeneity will be to divide C into two regions, in one of which $\|x\|$ is close to its maximum σ, and to concentrate on the separate entropies of these regions. In particular, from experience with Gaussian processes on R^1 (e.g., Berman (1985b)) we should expect that the distribution of $P\{\sup_C |Lx| > \lambda\}$ for large should be determined primarily by the entropy $N_C(\delta,\sigma,\varepsilon)$ as $\delta \nearrow \sigma$. This idea leads to the following result, in which, in most applications, we shall choose an f such that $f(\delta)\searrow 0$ as $\delta \nearrow \sigma$.

<u>Theorem 3.2</u> <u>Let</u> $f:(0,\sigma) \to \mathbb{R}$ <u>be such that there exist positive constants</u> a, κ <u>and</u> ε_0 <u>such that for all</u> $\varepsilon \in (0,\varepsilon_0]$, $\delta \in (0,\sigma)$,

$$(3.4) \qquad N_C(0,\delta,\varepsilon) \le a\varepsilon^{-\kappa}, \quad N_C(\delta,\sigma,\varepsilon f(\delta)) \le a\varepsilon^{-\kappa}.$$

<u>Then for each</u> δ <u>and all</u> $\lambda > \lambda^*(\varepsilon_0,\delta,\sigma,\kappa,f)$ <u>we have</u>

$$(3.5) \qquad P\{\sup_{x\in C} |Lx| > \lambda\}$$

$$\le \frac{5}{2} a(\sigma+1)\exp\{\frac{2(\sigma+\lambda^{-2})}{\sigma^4}\}\lambda^{-1}\{\lambda^{2\kappa}f^{\kappa}(\delta) + [\lambda^{-2} + \frac{(\sigma-\delta)}{2}]^{-\kappa}\}e^{-\lambda^2/2\sigma^2}$$

$$\le M\lambda^{-1}e^{-\lambda^2/2\sigma^2}\{\lambda^{2\kappa}f^{\kappa}(\delta) + [\lambda^{-2} + \tfrac{1}{2}(\sigma-\delta)]^{-\kappa}\}$$

<u>where</u> $M = \frac{5}{2} a(\sigma+\frac{1}{2})\exp\{\frac{2(\sigma+1)}{\sigma^4}\}$ <u>and</u> λ^* <u>is the smallest</u> λ <u>satisfying the following three conditions</u>:

$$(3.6) \qquad \lambda \ge [\min(\tfrac{1}{4},\varepsilon_0) - \frac{(\sigma-\delta)}{2}]^{-\frac{1}{2}},$$

(3.7) $\lambda \geq \max(2,\varepsilon_0^{-\frac{1}{2}}).f^{-\frac{1}{2}}(\delta)$,

(3.8) $$\lambda \geq \begin{cases} 2(\sigma + \tfrac{1}{2})^2(2\kappa+1) & 0 < \kappa < 4, \\ 2(\sigma + \tfrac{1}{2})^2(1+2\sqrt{\kappa}\,\ell n\kappa) & \kappa \geqslant 4 . \end{cases}$$

Note how the conditions on the constants are becoming unwieldy. To see how this result works, let us prove a simple corollary. The idea of the corollary is to introduce a parameter of "non-homogeneity", α, for C that describes the sizes of subsets of C over which $||x||$ is close to its overall supremum σ. Homogeneity is described by $\alpha = 0$, with increasing α describing increasing non-homogeneity. The result is

Corollary 3.1 Under the conditions of Theorem 3.2, if f satisfies

(3.9) $f(\delta) \leq c(\sigma-\delta)^{\alpha}$

for some positive α and c then for sufficiently large λ

(3.10) $P\{\sup_{x\epsilon C}|Lx| > \lambda\} \leq M\lambda^{-1 + 2\kappa/(1+\alpha)}e^{-\lambda^2/2\sigma^2}$,

where

$$M = \frac{5}{2}a(c + 2^{\kappa})(\sigma + \tfrac{1}{2})\exp 2(\sigma+1)/\sigma^4\} .$$

(The interested reader can easily substitute into (3.6) - (3.8) to make the statement "sufficiently large λ" more precise.)

Proof. Set $\delta = \sigma - \lambda^{-2/(1+\alpha)}$, taking λ large enough for δ to be positive. It is then straightforward to check that (3.6) - (3.8) are satisfied for large enough λ. Clearly, as $\lambda\to\infty$ we have $\delta\to\sigma$. To prove the corollary consider the last term in (3.5)

$$\lambda^{2\kappa} f^{\kappa}(\delta) + [\lambda^{-2} + \tfrac{1}{2}(\sigma-\delta)]^{-\kappa} \le c\lambda^{2\kappa}\lambda^{-2\kappa\alpha/(1+\alpha)} + (\lambda^{-2} + \tfrac{1}{2}\lambda^{-2/(1+\alpha)})^{-\kappa}$$

$$\le (c + 2^{\kappa})\lambda^{2\kappa/(1+\alpha)} ,$$

again for sufficiently large λ . Substituting this into (3.5) establishes the corollary.

Note, again, that large λ sends δ to σ. That is, it is only the neighborhood in C for which $\|x\|$ is close to σ that has any effect on the distribution of $\sup|Lx|$. To convince ourselves that the assumption (3.9) has actually led to a sharper bound, we need only note that the power of λ in (3.10) is never larger than that in (3.3), where no such assumption was made.

Our next assumption on C will be that it possesses some sort of scaling property, in the sense that there are subsets of C which look much like C itself, except that the original norm has been changed by a scaling factor. The idea then is to partition C into a number of smaller pieces, study the supremum on each one of these via Theorem 3.2, (to yield Theorem 3.3) and then piece the various bounds together to bound the supremum over C itself, (Theorems 3.4, 3.5). To this end, fix $\theta > 0$ and let G_θ be a partition of C satisfying

$$(3.11) \qquad \sup_{x,y \in A} \|x-y\| \le \theta \qquad \text{for all } A \in G_\theta .$$

Define $N_C^{G}(\theta) := \# G_\theta$. Clearly $N_C^{G}(\theta) \ge N_C(\theta)$, since the latter entropy is related to an G_θ of minimal cardinality. In general however we shall want to choose G_θ so that both entropies are effectively the same. Now we introduce the "scaling hypothesis", by assuming the existence of a function f and a constant a such that

(3.12) $\qquad N_A(f(\theta)\varepsilon) \leq a\varepsilon^{-\kappa} \quad$ for all $A \in G_\theta$,

and small enough $\varepsilon, \theta > 0$. Such an f always exists. (Take $f \equiv 1$!) Clearly, however, for this partitioning procedure to have any value, we shall want $f(\theta) \searrow 0$ as $\theta \searrow 0$. Nevertheless, it is not necessary to assume this at this stage, and the bounds in Theorem 3.3 and its corollaries are correct for any f. If f does not decrease to zero, however, they are uninteresting.

Note that it would be nice to replace (3.12) with the more pleasing condition $N_A(f(\theta)\varepsilon) \leq N_C(\varepsilon)$ comparing entropies. However, such a condition turns out to be impractical in examples, since we generally do not have the precise form of $N_C(\varepsilon)$, but only its growth rate.

Note, also, that we can always take $N_C^G(\theta)$ to be non-increasing, and, given some f satisfying (3.12), its left continuous monotone (non-decreasing) rearrangement also satisfies (3.12). Thus, in what follows, we shall always take f left continuous. Consequently, fixing some $p \geq 2$, the function

$$g(\theta) := \theta + 2f(\theta)/p^2$$

can also be taken to be left continuous, so that its inverse

$$g^{-1}(\eta) := \sup \{\theta: g(\theta) \leq \eta\}$$

is well defined. We can now state the following result which is closely related to Théorème 2.1.1 of Weber (1978)in the case $||x|| = 1$. Our style of proof is completely different however.

<u>Theorem 3.3</u> <u>Suppose</u> $N_C(\varepsilon) < a\varepsilon^{-\kappa}$ <u>for</u> $\varepsilon\epsilon(0,\varepsilon_0]$, <u>and that, for all</u> $\theta\epsilon(0,\theta_0]$, G_θ <u>and</u> f <u>satisfy</u> (3.12). <u>Then for every</u> $p \geq \max(2,\varepsilon_0^{-\frac{1}{2}})$, <u>any</u> $A\epsilon G_\theta$, $\sigma_A := \sup_A ||x||$, <u>and all</u> $\lambda > g(\theta)(1+4\kappa \ln p)^{\frac{1}{2}}$

$$\text{(3.13)} \qquad P(\sup_{x\epsilon A} |Lx| > \lambda) \leq \psi([\lambda - g(\theta)(1+4\kappa \ln p)^{\frac{1}{2}}]/\sigma_A) + 4ap^{2\kappa}\psi(\lambda/g(\theta)) + 4ap^{2\kappa}\sigma_A\lambda^{-1}e^{-\lambda^2/2\sigma_A^2} \cdot \exp(\lambda^2 g^2(\theta)/2\sigma_A^4)\,.$$

There is an easy corollary to this theorem that is far more illuminating. For large enough λ, set

$$(3.14) \qquad \theta_\lambda = g^{-1}([\lambda^2(1+4\kappa \ell np)]^{-\frac{1}{2}})$$

and substitute into (3.13). Then apply the standard inequality $\psi(u) < \sqrt{2/\pi}\, u^{-1} e^{-\frac{1}{2}u^2}$, $u \geq 0$, to obtain

Corollary 3.2 Under the conditions of Theorem 3.3 we have, for all $\lambda > \max(1.1, \{g(\sqrt{2})(1+4\kappa \ell np)^{\frac{1}{2}}\}^{-1})$, $A \in G_{\theta_\lambda}$,

$$P\{\sup_{x\in A} |Lx| > \lambda\} \leq c_1 \lambda^{-1} e^{-\frac{1}{2}\lambda^2/\sigma_A^2} + c_2 \lambda^{-2} \exp\{-\tfrac{1}{2}\lambda^4(1+4\kappa \ell np)\},$$

where

$$c_1 = 6\sigma_A e^{1/\sigma^2} + 4ap^{2\kappa}\sigma_A\ 2\exp\{(2\sigma_A^4(1+4\kappa \ell np))^{-1}\}$$

$$c_2 = 4ap^{2\kappa}(1+4\kappa \ell np)^{-\frac{1}{2}}$$

(The constant 6 in c_1 comes from $\lambda > 1.1$. In general, 6 can be replaced by $(1-\lambda^{-2})^{-1}$.)

An irritating aspect of both Theorem 3.3 and its corollary are that the constants diverge as $\sigma_A \to 0$. The same phenomenon occurs in Berman's (1985a) Theorem 3.1. In the following corollary, we show that this can easily be avoided via a simple trick, due, a referee tells us, to Lévy.

Corollary 3.3 Both Theorem 3.3 and Corollary 3.2 hold if we replace σ_A in the bounds by any $\sigma \geq \sigma_A$, as long as we then double the constants.

The proof is easy, so we give it now. Note firstly that if Z_t, $t \in T$, is any collection of a.s. bounded, zero mean, Gaussian variables, and Y an independent zero mean Gaussian variable, then

$$(3.15) \qquad P(\sup_t |Z_t| > \lambda) \le P(\sup_t Z_t > \lambda) + P(\inf_t Z_t < -\lambda)$$

$$= 2P(\sup_t Z_t > \lambda,\ Y \ge 0) + 2P(\inf_t Z_t < -\lambda,\ Y \le 0)$$

$$\le 2P(\sup_t |Z_t + Y| > \lambda)$$

To use this inequality, take $\sigma \ge \sigma_A$ and Y zero mean Gaussian with variance $\sigma^2 - \sigma^2{}_A$, independent of Lx for all $x \varepsilon A$, and define a new process L^* by $L^*x = Lx + Y$. Consider the image of A under L^*, call it A^*, as part of an L^2 space of Gaussian variables, where for any two points, u,v in the image such that $u = L^*x$, $v = L^*y$, $x,y \varepsilon A$ their inner product $(u,v)_*$ is given by $E(L^*x, L^*y)$. Then clearly

$$||u||_* = ||x|| + \sigma^2 - \sigma^2{}_A, \quad ||u-v||_* = ||x-y|| .$$

Consequently, $\sup_{A^*} ||u||_* = \sigma^2$ and A^* has the same entropy function as A. Let I be the identity map on this set. Then I is clearly isonormal on A^*, and $\sup_{A^*}|Iu| = \sup_A |L^*x|$. Thus, we can apply Theorem 3.3 and Corollary 3.2 to I and then note (3.15) with $Z = L$ to prove the corollary.

Now let us pause for a moment to consider the import of Theorem 3.3 and its corollaries. It is clear from Corollary 3.2 that for large λ, we find that the dominant term in the bound is $O(\lambda^{-1} e^{-\frac{1}{2}\lambda^2/\sigma_A{}^2})$. But this is of the order of the probability that a single zero mean Gaussian variable with variance $\sigma_A{}^2$ is greater than λ. That is, we have replaced the supremum of L over A by its value at one point only. Essentially, this has been done by making A small as λ becomes large, since $A \varepsilon G_{\theta_\lambda}$ and θ_λ will be small for λ large. That is, we have achieved at this stage a discretization of the supremum problem. This is actually the heart of the solution, for all we need do now is sum the bounds of Theorem 3.3 and its corollaries over the various sets in G_θ to bound the supremum over the whole of C.

To sum these bounds efficiently, we require further assumptions on the structure of C, as in the following two results, with which we complete this section, and in which we finally give up trying to keep track of constants. In the first result we shall, as in Theorem 3.2, concern ourselves primarily with regions of C of large norm.

Theorem 3.4 Suppose $N_C(\varepsilon) \le a\varepsilon^{-\kappa}$ for $\varepsilon \in (0,\varepsilon_0]$, and that there exist constants c and β such that for each $\theta \in (0,\theta_0]$ there exists a partition G_θ of C and constants $n_\theta, \delta_0(\theta)$ so that

$$(3.16) \qquad n(\delta,\theta) \le c(\sigma-\delta)^\beta N_C^{G}(\theta) + n_\theta \quad \text{for all } \delta \in (0,\delta_0(\theta)].$$

where

$$(3.17) \qquad n(\delta,\theta) := \#\{A \in G_\theta : A \cap C_\delta^+ \neq \phi\}.$$

Then, there exist constants c_1 and c_2 such that for sufficiently large λ

$$(3.18) \qquad P(\sup_{x\in C} |Lx| > \lambda) \le c_1 N_C^{G}(\theta_\lambda)\lambda^{-1-2\beta}(\ln\lambda)^\beta e^{-\lambda^2/2\sigma^2}$$

$$+ c_2 n_{\theta_\lambda} \lambda^{-1} e^{-\lambda^2/2\sigma^2} .$$

Here c_1 and c_2 depend on c,β,σ,δ_0, and an arbitrary p, but not on λ. The factor θ_λ is defined at (3.14).

Our final task is to free ourselves of the logarithmic term in (3.18) by partitioning C even more finely.

Theorem 3.5 Assume the assumptions of Theorem 3.4, but replace (3.16) by: There exists a $\Delta_0(\theta)$ such that for all $0 < \delta_2 - \delta_1 < \Delta_0$

$$n(\delta_1,\delta_2,\theta) \le c(\delta_2-\delta_1)^{\beta} N_C{}^{G}(\theta) + n_\theta, \tag{3.19}$$

where

$$n(\delta_1,\delta_2,\theta): = \#\{A\in G_\theta:\ A\cap C_{\delta_1}^{+}\cap C_{\delta_2}^{-} \neq \phi\}. \tag{3.20}$$

Then there exist constants C_1 and C_2 such that for sufficiently large λ

$$P(\sup_{x\in C} |Lx| > \lambda) \le c_1 N_C{}^{G}(\theta_\lambda)\lambda^{-1-2\beta} e^{-\lambda^2/2\sigma^2} + c_2 n_{\theta_\lambda} \lambda^{-1} e^{-\lambda^2/2\sigma^2}. \tag{3.21}$$

We shall now see how to apply these results to specific examples.

4. EXAMPLES

Our examples are of two kinds. In some we simply re-derive known results. Our aim here is to show that the rather general theorems of the previous sections give, when applied to specific cases, the best possible results. The more interesting examples which (by "induction") we also feel give the best possible bounds, are new. In particular, Examples 4.3 and 4.4, which consider the suprema of rectangle and half-plane indexed Brownian sheets, represent the first time sharp (asymptotic) bounds have been obtained for set indexed processes.

All our examples deal not with the isonormal process on Hilbert space H but with processes whose parameter space is generally somewhat simpler. Thus we shall have to translate these processes to the isonormal case. But this is easy, for if X_t is a Gaussian process on, say, a metric space (S,d) with continuous covariance function $R(s,t)$, then we simply identify H with the L^2 space of X, and $C \subset H$ with the set $\{x \in H\colon x = X_t$ form some $t \in S\}$. For $x=X_t$, $y=X_s$ in C we have $(x,y)_H = R(t,s)$. Clearly L is now the identity operator, so that Lx is simply x identified as a Gaussian variable rather than an element of H. Furthermore $\sup_{x \in C} |Lx| \equiv \sup_{t \in S} |X_t|$.

Entropy calculations are only slightly more involved, for we shall generally partition C by first partitioning S (this is usually geometrically simpler) and then letting the above identification induce a corresponding partition on C. We shall work the first example carefully to explain what is happening. In the later examples, we shall skimp on detail.

Example 4.1. Let X be a stationary, separable process on $[0,1]$ with zero mean and covariance function $R(t)$, which, for some positive a_1, β and γ_1 satisfies

$$(4.1) \quad 1 > R(t) > 1-a_1 t^{\beta} \qquad \text{for all } t \varepsilon [0,\gamma_1]$$

Let $\sigma(t)$ be a positive, continuous, monotonically increasing function on $[0,1]$ such that for some $\gamma_2 > 0$, $0 < a_2 \leq a_3$ and some $\alpha > 0$

$$(4.2) \qquad a_2|t-s|^{\alpha} \leq |\sigma(t) - \sigma(s)| \leq a_3|t-s|^{\alpha} \quad \text{whenever} \quad |t-s| < \gamma_2 .$$

Define now a scaled version of X by

$$Y(t) = \sigma(t)X(t), \qquad t \varepsilon [0,1].$$

We think of Y as a locally stationary process, (c.f. Berman (1974)) and shall show that

$$(4.3) \qquad P\{\sup_{[0,1]}|Y(t)| > \lambda\} \leq \begin{cases} c_1\lambda^{-1}e^{-\lambda^2/2\sigma^2(1)}, & \beta > \alpha > 0, \\ c_2\lambda^{-1-2/\alpha+2/\beta}e^{-\lambda^2/2\sigma^2(1)} & 0 < \beta \leq \alpha, \end{cases}$$

for some finite c_1 and c_2 and all $\lambda > 0$.

Before we prove (4.3), which we shall do via Theorem 3.5, it is instructive to consider how close we could get to (4.3) via existing theory. If we apply Berman's (1985a) recent bound, then the best we can do is a bound of the form

$$(4.4) \qquad P\{\sup_{[0,1]}|Y(t)| > \lambda\} \leq \begin{cases} c\lambda^{-1+1/\alpha}e^{-\lambda^2/2\sigma^2(1)} & \beta > 2\alpha > 0, \\ c\lambda^{-1+2/\beta}e^{-\lambda^2/2\sigma^2(1)} & 0 < \beta < 2\alpha. \end{cases}$$

This is clearly poorer than (4.3). [A proof of (4.4) follows easily from (4.5) below and Example 4.1 of Berman (1985a).] The above result could also be obtained, within the framework of this paper, via Theorem 3.3, which is effectively the analogue of Berman's result for the isonormal process.

One could also try to apply Weber's (1980) Théorème 2.1 here. In fact, his result is not strictly applicable, unless strict equality hold in (4.1) and (4.2). Assuming this, one obtains a result like (4.4), but with an extra factor of $\log \lambda$ in the bounds. Thus Weber's result is weaker yet than Berman's.

Finally, before commencing the proof, we note that bounds similar to (4.3) have been obtained for processes displaying covariance behaviour similar to that displayed by our $Y(t)$ by Piterbarg and Prisjaznjuk (1979). They actually do better than (4.3) for their case, for using arguments in the style of Pickands (1969a,b) they both identify the constants in their bound and show that the bound is sharp.

Throughout the proof we shall consider $Y(t)$ to be both a random variable and a point in H. From (4.1) and (4.2) we have that for all s,t with $|t-s| < \gamma_1 \wedge \gamma_2$

$$(4.5) \qquad \|Y(t) - Y(s)\|^2 = E(|\sigma(t)X(t) - \sigma(s)X(s)|^2) \le a_3^2|t-s|^{2\alpha} + 2\sigma^2(1)a_1|t-s|^{\beta} .$$

We now divide the argument to two distinct cases, and consider firstly $\beta \ge 2\alpha$. Then, via (4.5),

$$(4.6) \qquad \|Y(t) - Y(s)\| \le (a_3^2 + 2\sigma^2(1)a_1)^{\frac{1}{2}}|t-s|^{\alpha} := a_4|t-s|^{\alpha}.$$

To partition C, for each $\varepsilon > 0$, we simply partition the unit interval in sub-intervals each of length $(2\varepsilon/a_4)^{1/\alpha}$, and then map these intervals into C by the correspondence $t \to Y(t)$. Clearly then $N_C(\varepsilon) \leq (2\varepsilon/a_4)^{-1/\alpha}$ for small enough ε, so that we have polynomial entropy with $\kappa=1/\alpha$. (Actually, it is not quite true that $N_C(\varepsilon) \leq (2\varepsilon/a_4)^{-1/\alpha}$, for a true upper bound is $1 + [(2\varepsilon/a_4)^{-1/\alpha}]$, where, here, [x] is the integer part of x. Nevertheless, to make life a little easier, let us agree here that henceforth every time we bound an entropy by some non-integer, we allow ourselves the freedom of adding a minor "integer-correction factor", if necessary. This involves no real loss of precision.)

To obtain (4.3), we shall apply Theorem 3.5. For this we need a handle on the function $n(\delta_1,\delta_2,\theta)$ of (3.19), and to determine the θ_λ for this problem. To do this, fix θ, and let G_θ be the partition just described, but based on intervals of length $(\theta/a_4)^{1/\alpha}$. Subdividing each $A \varepsilon G_\theta$ according to the same principle, we easily obtain $N_C^G(\theta) = (\theta/a_4)^{-1/\alpha}$ and $N_A(\varepsilon\theta) \leq \varepsilon^{-1/\alpha}$ for small enough ε and θ. Fix $p \geq 2$, and compare this with (3.12). We see we can take $f(\theta) = \theta$ there, so that the $g(\theta)$ of (3.13) is given by $g(\theta) = \theta(1+2p^{-2})$, and the θ_λ of (3.14) by

$$(4.7) \qquad \theta_\lambda = \lambda^{-1}[(1+2p^{-2})(1+4\ell n p/\alpha)^{\frac{1}{2}}]^{-1}.$$

Now take $\sigma(0) \leq \delta_1 < \delta_2 \leq \sigma(1)$ and consider the set $C_{\delta_1}^+ \cap C_{\delta_2}^-$. It is easy to see (we leave the algebra to the reader) that for $\delta_2 - \delta_1 < a_3\gamma_2$ this set is the image of an interval in [0,1] of length between $a_3^{-1}(\delta_2-\delta_1)^{1/\alpha}$ and $a_2^{-1}(\delta_2-\delta_1)^{1/\alpha}$.

To finally bound $n(\delta_1,\delta_2,\theta) := \#\{A \varepsilon G_\theta:\ A \cap C_{\delta_1}^+ \cap C_\delta^- \neq \emptyset\}$ for

$|\delta_2-\delta_1| < \Delta_0(\theta)$, set $\Delta_0(\theta) = \theta^2$. Then since each $A\varepsilon G_\theta$ is the image of an interval of length $O(\theta^{1/\alpha})$ and $C_{\delta_1}^+ \cap C_{\delta_2}^-$ the image of an interval of length at most $O(\theta^{2/\alpha})$, we have that $n(\delta_1,\delta_2,\theta)$ is at most two. Thus (3.19) is satisfied for all positive β with $n_\theta = 2$, and all the conditions of Theorem 3.5 are satisfied. Consequently (3.21) holds for every $\beta \geq 2\alpha$. Take β arbitrarily large in (3.21) to obtain (4.3) and prove our result for the case $\beta \geq 2\alpha$.

Now take $\beta < 2\alpha$. Then by (4.6) we have for small $|t-s|$ tnat $||Y(t) - Y(s)|| \leq a_4 \; |t-s|^{\beta/2}$. The argument above thus gives colynomial entropy with $\kappa = 2/\beta$. Defining G_θ as the image of intervals of length $(\theta/a_4)^{2/\beta}$, we once again find $f(\theta) = \theta$ but now

$$\theta_\lambda = \lambda^{-1}[(1+2p^{-2})(1+ \tfrac{8}{\beta}\ell np)^{\frac{1}{2}}]^{-1} .$$

The set $C_{\delta_1}^+ \cap C_{\delta_2}^-$ remains as it was above. Again take $\Delta_0(\theta) = \theta^2$ and consider $n(\delta_2,\delta_1,\theta)$. If $\beta \varepsilon (\alpha,2\alpha)$ then, once again, as $\theta \searrow 0$ we find $n(\delta_1,\delta_2,\theta) \leq 2$ for $\delta_2 - \delta_1 < \Delta_0(\theta)$. Consequently, in this case the argument is precisely as above, and we now have (4.3) for all $\beta > \alpha$.

If $0 < \beta \leq \alpha$ the intervals mapping into G_θ can be shorter than those mapping onto the $C_{\delta_1}^+ \cap C_{\delta_2}^-$, (lengths $O(\theta^{2/\beta})$ versus $(\delta_2 - \delta_1)^{1/\alpha} \leq O(\theta^{1/\alpha})$). Consequently, noting that $N_C^{\,G}(\theta) = (\theta/a_4)^{-2/\beta}$, we obtain

$$n(\delta_1,\delta_2,\theta) \leq 2 + c(\delta_2 - \delta_1)^{1/\alpha}N_C^{\,G}(\theta)$$

for some finite c. That is, we have the right bound for (3.19) of Theorem 3.5. Substitution into (3.21) completes the proof.

Our remaining examples are all connected with Brownian sheets. Let λ_k be Lebesque measure on $[0,1]^k$. The zero mean Gaussian process W defined on Borel sets in $[0,1]^k$ with covariance

$$(4.8) \qquad E[W(A)W(B)] = \lambda_k(A \cap B),$$

is called the set indexed Brownian sheet. The pinned version of W, denoted by

$$\overset{\circ}{W}(A) := W(A) - \lambda_k(A)W([0,1]^k)$$

has covariance

$$(4.9) \qquad E(\overset{\circ}{W}(A)\overset{\circ}{W}(B)) = \lambda_k(A \cap B) - \lambda_k(A)\lambda_k(B).$$

For the special case of W indexed only by k-intervals of the form $A_{\underset{\sim}{t}} = \prod_{i=1}^{k} [0,t_i]$, we write $W(\underset{\sim}{t}) := W(A_{\underset{\sim}{t}})$ and $\overset{\circ}{W}(\underset{\sim}{t}) = \overset{\circ}{W}(A_{\underset{\sim}{t}})$, and call $W(\underset{\sim}{t})$ and $\overset{\circ}{W}(\underset{\sim}{t})$ the point indexed sheet and pinned sheet, respectively.

$W(\underset{\sim}{t})$ is of particular interest as the natural k-dimensional generalisation of Brownian motion while $\overset{\circ}{W}(A)$ arises as a weak limit in an empirical measure setting. (c.f. Dudley (1978).) We start with the point indexed pinned sheet.

Example 4.2 Let $\overset{\circ}{W}$ be a point indexed Brownian sheet on $[0,1]^k$. Then there exists a finite c such that

$$(4.10) \qquad P\{\sup_{[0,1]^k} |\overset{\circ}{W}(\underset{\sim}{t})| > \lambda\} \leq c\lambda^{2(k-1)}e^{-2\lambda^2}.$$

This result was originally established in somewhat greater generality in Adler and Brown (1985), where it was also shown that this bound serves, for different c, as a lower bound as well. It is not, however, obtainable from any other general Gaussian bound. Using Berman's (1984a) result, or our Theorem 3.3, the best bound possible is only $O(\lambda^{2k-1}e^{-2\lambda^2})$.

We rederive the result here to show how it can be obtained from the general theory. Once again, we shall apply Theorem 3.5, so we are basically concerned with finding a good bound for $n(\delta_1,\delta_2,\theta)$, and the other factors in (3.19).

We commence by noting

$$(4.11) \qquad \|\mathring{W}(\underset{\sim}{t}) - \mathring{W}(\underset{\sim}{s})\|^2 = E[(\mathring{W}(\underset{\sim}{t}) - \mathring{W}(\underset{\sim}{s}))^2]$$

$$\leq \lambda(A_{\underset{\sim}{t}} \Delta A_{\underset{\sim}{s}}) \leq \sum_{i=1}^{k} |t_i - s_i| ,$$

for all $\underset{\sim}{s},\underset{\sim}{t} \in [0,1]^k$. Now, for each $\theta > 0$ set $m_\theta := [k\theta^{-2}]$ ($[x]: =$ integer part of x) and define the partition I_θ of $[0,1]^k$ by

$$I_\theta = \{A \subset [0,1]^k : A = \prod_{i=1}^{k} (\frac{n_i}{m_\theta}, \frac{n_i+1}{m_\theta}] \; n_i = 0,1,\ldots,m_\theta - 1\} .$$

Furthermore, let G_θ be the partition I_θ induces in H, the L^2 space of $\mathring{W}$. By (4.11), if $x,y \in A \in G_\theta$, then $\|x-y\| \leq \theta$, so that G_θ is a partition of the type required for Theorem 3.5, and

$$(4.12) \qquad N_C^{G}(\theta) = [k/\theta^2]^k \leq 3k^k \theta^{-2k} ,$$

the inequality following by simple algebra. By (4.12), C has polynomial entropy with $\kappa \leq 2k$. We now check the scaling property.

Fix $\varepsilon > 0$, set $p_\varepsilon := [\varepsilon^{-2}]$, divide each $A \in I_\theta$ into $p_\varepsilon{}^k$ equal k-intervals, and map these into the corresponding $A \in G_\theta$. Applying (4.11) once again, it is easy to check that

$$N_A(\theta\varepsilon) \leq 3\varepsilon^{-2k} \quad \text{for all } \varepsilon < (2k)^{-\frac{1}{2}} \text{ and } A \in G_\theta .$$

Thus we can take $f(\theta) = \theta$ in (3.12) and, for some $p \geq 2$,

$$\text{(4.13)} \qquad \theta_\lambda = \lambda^{-1}[(1+2p^{-2})(1+8k\ell np)^{\frac{1}{2}}]^{-1} .$$

All that remains is to investigate $n(\delta_1,\delta_2,\theta)$. Firstly note that it suffices to consider $\delta_1 > \frac{1}{4}$, for we can break up C into two parts, over which $\|x\| \leq \frac{1}{4}$ and $\|x\| > \frac{1}{4}$. Over the first part the inequality (1.1) gives us an upper bound of $O(e^{-8\lambda^2})$ for the tail of the supremum, which is clearly of smaller order than the desired (4.10). Thus the case $\delta_i \leq \frac{1}{4}$ can be neglected. Now note that $C^+_{\delta_1} \cap C^-_{\delta_2}$ is the image of the following set, in which we write $|\underset{\sim}{t}|$ for $t_1 \times \ldots \times t_k$.

$$\text{(4.14)} \qquad I(\delta_1,\delta_2) = \{\underset{\sim}{t}\colon \delta_1{}^2 \leq |\underset{\sim}{t}|(1-|\underset{\sim}{t}|) \leq \delta_2{}^2\}$$

$$= \{\underset{\sim}{t}\colon \tfrac{1}{2} - (\tfrac{1}{4}-\delta_1{}^2)^{\frac{1}{2}} \leq |\underset{\sim}{t}| \leq \tfrac{1}{2} - (\tfrac{1}{4}-\delta_2{}^2)^{\frac{1}{2}}\}$$

$$\cup\{\underset{\sim}{t}\colon \tfrac{1}{2} + (\tfrac{1}{4} - \delta_2{}^2)^{\frac{1}{2}} \leq |\underset{\sim}{t}| \leq \tfrac{1}{2} + (\tfrac{1}{4} - \delta_1{}^2)^{\frac{1}{2}}\}$$

The second line follows via a little elementary algebra. To count the number of A from I_θ that intersect $I(\delta_1, \delta_2)$ it suffices to count the number of lattice points of the form $(n_1/m_\theta,\ldots,n_k/m_\theta)$ falling in $I(\delta_1,\delta_2)$. But this is relatively easy, for if we fix $n_1,\ldots,n_{k-1}$ then some more algebra applied to (4.14) shows that no more than $32\sqrt{2}(\delta_2-\delta_1)^{\frac{1}{2}}m_\theta$ values of n_k are permissible. Allowing $n_1,\ldots,n_{k-1}$ to vary, we thus obtain

$$\begin{aligned} n(\delta_1,\delta_2,\theta) &\le c(m_\theta)^{k-1}(\delta_2-\delta_1)^{\frac{1}{2}}m_\theta \\ &\le c(k)\theta^{-2k}(\delta_2-\delta_1)^{\frac{1}{2}} \\ &\le c(\delta_2-\delta_1)^{\frac{1}{2}}N_C^{G}(\theta) . \end{aligned}$$

But this is all we need, for substitution into (3.21), on noting that $\sigma^2 = \frac{1}{4}$ for this problem, immediately establishes the required (4.10).

Example 4.3 Let R_k be the set of all k-intervals of the form $[\underset{\sim}{s},\underset{\sim}{t}] = \prod_{i=1}^{k} [s_i,t_i]$ contained in $[0,1]^k$. Then there exists a constant c such that

$$(4.15) \qquad P\{\sup_{R_k}|\overset{\circ}{W}(A)|>\lambda\} \le c\lambda^{2(2k-1)}e^{-2\lambda^2}.$$

Before we prove this result, we shall establish its sharpness by showing that there exists a c' such that

$$(4.16) \qquad c'\lambda^{2(2k-1)}e^{-2\lambda^2} \le P\{\sup_{R_k} \overset{\circ}{W}(\lambda) > \lambda\}.$$

We shall prove this for k=2. For k>2 the proof is basically the same, the notation is just a little longer. Let $A = [\underset{\sim}{s},\underset{\sim}{t}]$ be a rectangle in $[0,1]^2$, and define a mapping $T:R_2 \to [0,1]^4$ by

$$T([\underset{\sim}{s},\underset{\sim}{t}]): = \left(\frac{t_1-s_1}{t_1},\ t_1, \frac{t_2-s_2}{t_2},\ t_2\right).$$

Clearly we must have $0 \leq s_i \leq t_i \leq 1$, $i=1,2$ for $[\underset{\sim}{s},\underset{\sim}{t}]$ to be in R_2, and so it is easy to see that T is one-one and onto. The inverse mapping is defined by

$$(4.17)\qquad T^{-1}(z_1,z_2,z_3,z_4) = [(z_2(1-z_1),z_4(1-z_3)),\ (z_2,z_4)].$$

Now define a process $X(\underset{\sim}{z})$ on $[0,1]^4$ by $X(\underset{\sim}{z}) = \overset{\circ}{W}(T^{-1}(\underset{\sim}{z}))$. This process is clearly Gaussian with zero mean, and it follows from (4.17) and (4.8) that

$$(4.18)\qquad E[X^2(\underset{\sim}{z})] = \lambda(T^{-1}(\underset{\sim}{z})) = |\underset{\sim}{z}| - |\underset{\sim}{z}|^2$$

This is the variance of the point indexed sheet on $[0,1]^4$. After a page or so of elementary algebra, one can also derive the rather useful inequality that for any $A,B \in R_2$,

$$\lambda(A \cap B) \leq \prod_{i=1}^{4} [T_i(A) \wedge T_i(B)]$$

where $T_i(A)$ is the i-th coordinate of $T(A)$. An immediate consequence of this is that

$$E[X(\underset{\sim}{u})X(\underset{\sim}{v})] = E[\overset{\circ}{W}(T^{-1}\underset{\sim}{u})\overset{\circ}{W}(T^{-1}\underset{\sim}{v})] \leq \prod_{i=1}^{4} u_i \wedge v_i - |\underset{\sim}{u}|\cdot|\underset{\sim}{v}|.$$

That is, the covariance function of X is dominated by that of the point indexed sheet on $[0,1]^4$. Consequently, by (4.18) and Slepian's inequality (Slepian (1962)), the tail of supX dominates that of the sheet. Theorem 2.2 of Adler and Brown (1985) states that this, in turn dominates $c'\lambda^6 e^{-2\lambda^2}$ for some c', (or $c'\lambda^{2(2k-1)}e^{-2\lambda^2}$ for general k), so that (4.16) is proven.

Now to the upper bound. We shall give the main steps of the derivation and skip all the algebra, most of which is similar to that in the previous example. To define G_θ , set $m_\theta = [2k/\theta^2]$, and let G_θ be the image in H of the partition of R_k given by $\cup_{\underset{\sim}{J}\in L_k(\theta)} A(\underset{\sim}{J})$ where $L_k(\theta)$ is the set of all integer 2k-tuples of the form $(j_1^{(1)}, j_1^{(2)}, \ldots, j_k^{(1)}, j_k^{(2)})$ with $j_i^{(1)} \leq j_i^{(2)}, \ldots, i=1,\ldots,k$, $j_i^{(\ell)} = 0,1,\ldots,m_\theta-1$, $i=1,\ldots,k$, $\ell=1,2$ and $A(\underset{\sim}{J})$ is the collection of all k-intervals $[\underset{\sim}{x},\underset{\sim}{y}]$ satisfying $|x_i - j_i^{(1)}/m_\theta| < \theta^2/2k$, $|y_i - j_i^{(2)}/m_\theta| < \theta^2/2k$, $i = 1,\ldots,k$. It is easy to see that G_θ is a partition of the required form, and that

$$N_C^{G}(\theta) \leq 3.4^k k^{2k} \theta^{-4k} = c\theta^{-4k} .$$

Consequently we have polynomial entropy with parameter $\kappa = 4k$. Continuing the same procedure, it is easy to see that, for each $A \epsilon G_\theta$, $N_A(\theta\varepsilon) \leq c\varepsilon^{-4k}$, so that as in the previous case we have $f(\theta) = \theta$ and $\theta_\lambda = c\lambda^{-1}$.

Now consider $C^+_{\delta_1} \cap C^-_{\delta_2}$, which we can write as

$$\{ B = \prod_{i=1}^{k} [x_i,y_i]: \quad \delta_1^{\,2} \leq \prod_1^k (y_i - x_i) - [\prod_1^k (y_i - x_i)]^2 \leq \delta_2^{\,2} \} .$$

Again we can assume $\delta_1 > \frac{1}{4}$, and follow the procedure of the previous example to eventually obtain

$$n(\delta_1,\delta_2,\theta) \leq cN_C^{G}(\theta)(\delta_2-\delta_1)^{\frac{1}{2}} \qquad \text{for } \delta_2-\delta_1 < C\theta^2.$$

Substituting all the above into Theorem 3.5, together with the fact that $\sigma=\frac{1}{2}$, we prove (4.15)

The previous two examples almost seem to indicate that in working with Brownian sheets it is only the dimensionality, d, of the parameter space that determines the power of λ in our bound. For example, for $\overset{\circ}{W}$ on $[0,1]^k$, we have $d=k$ and the bound is $c\lambda^{2(d-1)}e^{-2\lambda^2}$. For $\overset{\circ}{W}$ on R_k, we have $d=2k$ (each $A \in R_k$ can be specified by $2k$ parameters) and the bound is again $C\lambda^{2(d-1)}e^{-2\lambda^2}$. We find this once again in treating $\overset{\circ}{W}$ indexed by all half-squares in R^2, (which we shall write as $\mathcal{D}_2 := \{A \subset [0,1]^2 : A = [0,1]^2 \cap \{(x,y): \alpha x + \beta y + \gamma \leq 0 \text{ some } \alpha,\beta,\gamma \in [-\infty,\infty]\}\}$, for which $d=2$.

<u>Example 4.4</u> For the Brownian sheet indexed by half-squares, we have

$$(4.19) \qquad P\{\sup_{\mathcal{D}_2} |\overset{\circ}{W}(A)| > \lambda\} \leq c\lambda^2 e^{-2\lambda^2},$$

for some finite, positive c.

To commence the proof of (4.19) note firstly that if $A \in \mathcal{D}_2$, then $\overset{\circ}{W}(A) = -\overset{\circ}{W}(A^c)$. Consequently we need only consider half of $\mathcal{D}_2$, say those half squares that contain at least one of the points $(1,0)$ or $(1,1)$. We write this as $\mathcal{D}_2^+$.

Let $S_1,\ldots,S_4$ denote the four sides of the unit square, $\{(x,y): 0 \leq x,y \leq 1\}$ on which, respectively, $x=0$, $x=1$, $y=0$, $y=1$. To define G_θ, set $m_\theta = [\theta^{-2}]$ and $x_i^{(k)}(\theta)$ the point on S_k at a distance i/m_θ from its start. Now let $A(\theta,k,\ell,i,j)$ be the collection of all half planes in $\mathcal{D}_2^+$ with boundary intersecting S_k between $x_i^{(k)}$ and $x_{i+1}^{(k)}$, and S_ℓ between $x_j^{(\ell)}$ and $x_{j+1}^{(\ell)}$. $(k,\ell=1,\ldots,4, k\neq\ell, i,j=0,1,\ldots,m_\theta^{-1})$. These A clearly provide a partition of $\mathcal{D}_2^+$, and we take the induced partition in the L^2 space of $\overset{\circ}{W}$ as G_θ. Clearly G_θ has the properties we generally require and, furthermore

$$(4.20) \qquad N_C^{\,G}(\theta) = \binom{4}{2}(m_\theta+1)^2 \le 24\theta^{-4} \ .$$

Consequently we have polynomial entropy with $\kappa=4$. To further subdivide these sets, simply subdivide each interval $[x_i^{(k)}, x_{i+1}^{(k)}]$ more finely, so that simple calculations yield that $N_A(\varepsilon\theta) \le 4\varepsilon^{-4}$ for each such A. Consequently $f(\theta) = \theta$ and for $p \ge 2$

$$\theta_\lambda = \lambda^{-1}[(1+2p^{-2})(1+8\ell np)^{\frac{1}{2}}]^{-1} \ .$$

It remains to estimate $n(\delta_1,\delta_2,\theta)$, for which we must describe $C^+_{\delta_1} \cap C^-_{\delta_2}$. As before, this is made up of the image of all half squares whose intersections with $[0,1]^2$ have area S satisfying either

$$(4.21) \qquad a_1 = \tfrac{1}{2} + (\tfrac{1}{4} - \delta_2^{\,2})^{\frac{1}{2}} \le S \le \tfrac{1}{2} + (\tfrac{1}{4} - \delta_1^{\,2})^{\frac{1}{2}} = b_1$$

or

$$(4.22) \qquad a_2 = \tfrac{1}{2} - (\tfrac{1}{4} - \delta_1^{\,2})^{\frac{1}{2}} \le S \le \tfrac{1}{2} - (\tfrac{1}{4} - \delta_2^{\,2})^{\frac{1}{2}} = b_2$$

We further divide $C^+_{\delta_1} \cap C^-_{\delta_2}$, into the image of half squares whose intersection with $[0,1]^2$ is a proper quadrilateral, and those that yield a triangle. We shall count only the first case, the second can be treated similarly, and yields same order of magnitude bounds on $n(\delta_1,\delta_2,\theta)$. Clearly, because of symmetry, we need only treat quadrilaterals including all of the side S_2, for we then simply add a factor of two to our counting to account for the side S_3.

Such quadrilaterals can be parametrized by two points u and v representing, respectively the points of intersection of the boundary of the half plane with the sides S_3 and S_4 of $[0,1]^2$. Then the area of the quadrilateral is given by $1-\frac{1}{2}(u+v)$. For such a quadrilateral to be in the pre-image of $C^+_{\delta_1} \cap C^-_{\delta_2}$ it thus follows from (4.21) and (4.22) that

$$2(1-b_i) \leq u + v \leq 2(1-a_i) \quad \text{for } i=1 \text{ or } 2.$$

Similarly, if the coordinates $x^{(3)}_{i_1}(\theta)$ and $x^{(4)}_{i_2}(\theta)$ on S_3 and S_4 define a half square whose image lies in $C^+_{\delta_1} \cap C^-_{\delta_2}$, then

$$(4.23) \qquad 2(1-b_i)m_\theta \leq i_1+i_2 \leq 2(1-a_i)m_\theta \quad \text{for } i=1 \text{ or } 2.$$

For fixed a_i, b_i the number of pairs (i_1, i_2) satisfying (4.23) is no more than $2m^2_\theta(b_i-a_i)$. Now note that via a little algebra

$$16(b_i-a_i) = (1-4\delta_1{}^2)^{\frac{1}{2}} - (1-4\delta_2{}^2)^{\frac{1}{2}} \leq c(\delta_2-\delta_1)^{\frac{1}{2}}$$

Using this and all the above we find that for small enough $\delta_2-\delta_1$,

$$n(\delta_1,\delta_2,\theta) \leq c(\delta_2-\delta_1)^{\frac{1}{2}}m_\theta{}^2$$

$$= c(\delta_2-\delta_1)^{\frac{1}{2}}N^G(\theta) .$$

Now apply Theorem 3.5 and the fact that $\sigma = \frac{1}{2}$ to obtain (4.19) and so complete the proof.

5. PROOFS FOR SECTION 3

We need firstly to establish (3.2), i.e. for $p \geq 2$ and all $\lambda > (1+4\kappa \ln p)^{\frac{1}{2}}$

$$(5.1) \qquad P\{\sup_C |Lx| > \lambda(\sigma+2p^{-2})\} \leq \frac{5}{2}\, a p^{2\kappa} \int_\lambda^\infty e^{-\frac{1}{2}u^2} du \, .$$

Our starting point is the basic inequality (2.9). There, put m=1, so that $\delta_0=0$, $\delta_1 = \sigma$, and there is only one λ sequence and one ε sequence. Set $\varepsilon_j = p^{-2^j}$ and $\lambda_j = \lambda 2^{j/2}$. Then (2.9) becomes

$$(5.2) \qquad P\{\sup_C |Lx| > \lambda(\sigma + \sum_{j=1}^{\infty} 2^{j/2} p^{-2^j})\} \leq a \sum_{j=0}^{\infty} p^{\kappa 2^{j+1}} \psi(\lambda 2^{j/2}) \, .$$

The sums are easy to calculate. Following Fernique (1975), for $j \geq 0$

$$p^{-\kappa^{j+1}} \psi(\lambda 2^{j/2}) = \sqrt{2/\pi} \int_\lambda^\infty \exp[\kappa 2^{j+1} \ln p + \tfrac{1}{2} j \ln 2 - u^2 2^{j-1}] du$$

$$\leq \sqrt{2/\pi} \int_\lambda^\infty \exp[-\tfrac{1}{2}u^2 + 2\kappa \ln p + \tfrac{1}{2}(j \ln 2 + 1-2^j)] du,$$

if $\lambda > (1+4\kappa \ln p)$. Consequently the rightmost sum in (5.2) is bounded by

$$a p^{2\kappa} \psi(\lambda) \sum_{j=0}^{\infty} 2^{j/2} \exp \tfrac{1}{2}(1-2^j).$$

Evaluating the sum gives the upper bound in (5.1) with a little room to spare for the constant. The leftmost sum in (5.2) is easily bounded by $2p^{-2}$, and so (5.1) is established.

We can now start proving the theorems of section 3.

<u>Proof of Theorem 3.1</u> We commence with (5.1). Note firstly from the proof of (5.1) we require $p^{-2} \leq \varepsilon_0$, i.e. $p > \varepsilon_0^{-\frac{1}{2}}$. We have also required $p \geq 2$.

Then by (5.1) and the fact $p \geq 2$ we have that for $\lambda > (\sigma+\frac{1}{2})^2(1+4\kappa \ln p)^{\frac{1}{2}}$

$$(5.3) \qquad P\{\sup_C |Lx| > \lambda\} \leq \frac{5}{2} ap^{2\kappa} \int_{\lambda/(\sigma+2p^{-2})}^{\infty} e^{-u^2/2} du$$

$$\leq \frac{5}{2} ap^{2\kappa}\lambda^{-1}(\sigma+2p^{-2})\exp\{-\lambda^2/2(\sigma+2p^{-2})^2\} ,$$

the last line via the standard inequality. Now set $p=\lambda$ in (5.3), which can be done if we take $\lambda \geq \max(2,\varepsilon_0^{-\frac{1}{2}})$ and $\lambda > (\sigma+\frac{1}{2})^2(1+4\kappa \ln\lambda)$. Simple algebra converts these to the conditions on λ given in the statement of the theorem. Then on substitution, we obtain

$$(5.4) \qquad P\{\sup_C |Lx| > \lambda\} \leq \frac{5}{2} a\lambda^{2\kappa-1}(\sigma+\tfrac{1}{2})\exp\{-\lambda^2/2(\sigma+2\lambda^{-2})^2\} .$$

Under the conditions we have on λ , it is easy to check that the exponent here is bounded above by $\lambda^2/2\sigma^2 - 2(\sigma+\lambda^{-2})/\sigma^4$. This completes the proof.

<u>Proof of Theorem 3.2</u> We shall not keep track of the constants of the Theorem throughout. Doing so more than doubles the length of, and complicates, an otherwise simple argument. The interested reader can check the constants by adding to the following argument some simple algebra.

Fix $\delta \in (0,\sigma)$, choose λ large, and note that we can always choose f so that $f(\delta) \leq 1$. Then define

$$p_1 = [\lambda^{-2}+\tfrac{1}{2}(\sigma-\delta)]^{-\frac{1}{2}} , \qquad p_2 = \lambda f^{\frac{1}{2}}(\delta) .$$

Both p_1 and p_2 are less than λ . Apply (5.3) to the two sets $C_1 := C^+_0 \cap C^-_\delta$ and $C_2 := C^+_\delta \cap C^-_\sigma$, using p_1 and p_2, respectively, in place of the p there. We find

$$(5.5)\qquad P(\sup_{C_1}|Lx| > \lambda) \le c[\lambda^{-2}+\tfrac{1}{2}(\sigma-\delta)]^{-\kappa}\lambda^{-1}e^{-\lambda^2/2\sigma^2},$$

bounding the exponent in (5.3) as in (5.4). Furthermore

$$(5.6)\qquad P(\sup_{C_2}|Lx| > \lambda) \le c[\lambda f^{\frac{1}{2}}(\lambda)]^{2\kappa}\lambda^{-1}e^{-\lambda^2/2\sigma^2} .$$

Combining (5.5) and (5.6) proves the theorem.

Proof of Theorem 3.3 The idea of the proof is simple. If θ is small, then so are the sets in G_θ. For $A \in G_\theta$, choose some $x^* \in A$. For each $x \in A$, write $L = Lx^* + L(x-x^*)$. Since $\|x-x^*\|$ must be small, $L(x-x^*)$ should be also small (stochastically). To show this we consider $L(x-x^*)$ conditional on Lx^*, using an idea used previously in Adler and Brown (1985) and Berman (1984a) for certain Gaussian processes on R^k. Consequently, $Lx = Lx^*$ + a smaller order term. Precise estimates are given in the theorem. The details of the proof are as follows.

Take $A \in G_\theta$ and let x^* be a point in A satisfying $\|x^*\| = \sup_A \|x\|$, i.e. x^* has maximal norm in A. (Such an x^* exists, for we lose no generality in assuming A closed, and our assumption of finite entropy then guarantees compactness and so the existence of x^*.) Consider the process

$$L^*x: = L(x-x^*) = Lx - Lx^* ,$$

and let A^* be its image in $L^2(\Omega,P)$. Let I be the (identity) operator on A^* that simply identifies each element of A^* as a Gaussian variable. The inner product (u,v) of $u=L^*x$ and $v=L^*y$ in A^* is given by

$E(L^*x.L^*y)$, I is isonormal on A^* and $\sup_{A^*}|Iu| = \sup_C|L^*x|$. Furthermore, it is trivial to check that

$$\sup_{A^*} \|u\|_* \le \theta^2, \quad \text{and} \quad \|u-v\|_* = \|x-y\|$$

Thus the entropy function for I is identical to that for L on the original space. Now recall the proof of (5.1). Rework it for I on A^*, noting condition (3.12), with p replaced by $pf^{-\frac{1}{2}}(\theta)$. This gives

$$(5.8) \qquad P\{\sup_A |L^*x| > \lambda[\theta + 2f(\theta)p^{-2}]\} \le \frac{5}{2} a p^{2\kappa} \int_\lambda^\infty e^{-u^2/2}\, du .$$

Furthermore, precisely the same bound holds if we replace L^*x by $L^{**}(x) := Lx - E(Lx|Lx^*)$. This follows as for L^*, on noting that $\|u-v\|_{**} \le \|u-v\|_*$, which follows from an easy calculation on conditional variances.

Now note that the event that interests us, $\sup_A |Lx| > \lambda$, is included in the union of the four events:

$$(5.9) \qquad |Lx^*| > \lambda - g(\theta)(1+4\kappa \ell n p)^{\frac{1}{2}},$$

$$(5.10) \qquad \sup_A |L^*x| > \lambda,$$

$$(5.11) \qquad \sup_A Lx > \lambda \quad \text{and} \quad 0 \le Lx^* \le \lambda - g(\theta)(1+4\kappa \ell n p)^{\frac{1}{2}},$$

$$(5.12) \qquad \inf_A Lx < -\lambda \quad \text{and} \quad -\lambda + g(\theta)(1+4\kappa \ell n p)^{\frac{1}{2}} \le Lx^* \le 0.$$

The probability of (5.9) is bounded by the first term in (3.13), while the second term there bounds the probability of (5.10) by (5.8).

The probabilities of (5.11) and (5.12), which are clearly identical, are a little more involved to derive.

Note first that by well known properties of Gaussian variables

$$E(Lx|Lx^* = \eta) = \frac{(x,x^*)}{(x^*,x^*)}\eta \le \eta$$

if $\eta > 0$, since x^* is a point of maximal norm. Consequently $E(Lx|Lx^*) \le Lx^*$ on the set where $Lx^* \ge 0$, and so (5.11) is contained in the event

$$\sup_A L^{**}x > \lambda - Lx^* \quad \text{and} \quad 0 \le Lx^* \le \lambda - g(\theta)(1+4\kappa \ell n p)^{\frac{1}{2}}.$$

But $L^{**}x$ and Lx^* are independent, so the probability of this event can be bounded by

$$\int_0^\gamma P(\sup_A L^{**}x > \lambda - u)p(u/\sigma_A)\sigma_A^{-1}du$$

with $\gamma = \lambda - g(\theta)(1+4\kappa \ell n p)^{\frac{1}{2}}$. Applying (5.8) for L^{**}, we can bound this by

$$2ap^{2\kappa}\int_0^\lambda \psi(\frac{\lambda - u}{g(\theta)})p(u/\sigma_A)\sigma_A^{-1}\ du\ .$$

Setting $z = \lambda(\lambda - u)$, this can be further bounded by

$$2ap^{2\kappa}\int_0^\infty \psi\ (\frac{z}{\lambda g(\theta)}\ )p\,(\frac{\lambda - z/\lambda}{\sigma_A})(\lambda\sigma_A)^{-1}dz\ .$$

Noting that $p(x+y) \le p(x)e^{-xy}$ for all x,y, we can further bound the above by

$$\frac{2ap^{2\kappa}}{\lambda\sigma_A} p(\frac{\lambda}{\sigma_A})\int_0^\infty \psi(\frac{z}{\lambda g(\theta)})e^{z/\sigma_A^2}\, dz \ .$$

This is now a standard integral, and turns out to be no more than half the last factor in (3.13). This completes the proof of Theorem 3.3.

<u>Proof of Theorem 3.4</u>. Consider Corollary 3.2 for A's belonging to $G_\theta \cap C_\delta^+$ and $G_\theta \cap C_\delta^-$, where $\delta\in(0,\sigma)$. Noting the dependence of c_1 in Corollary 3.2 on σ by writing $c_1(\sigma)$, we find

$$P(\sup_C|Lx| > \lambda) \leq n(\delta,\theta_\lambda)\{c_1(\sigma)\lambda^{-1}e^{-\lambda^2/2\sigma^2}+c_2\lambda^{-2}\exp[-\tfrac{1}{2}\lambda^4(1+4\kappa\ell np)]\}$$

$$+ [N^G(\theta_\lambda)-n(\delta,\theta_\lambda)]\cdot\{c_1(\delta)\lambda^{-1}e^{-\lambda^2/2\delta^2}+c_2\lambda^{-2}\exp[-\tfrac{1}{2}\lambda^4(1+4\kappa\ell np)]\}$$

Along with the other restrictions on λ, now take $\lambda > [\delta^2(1+4\kappa\ell np)]^{-\frac{1}{2}}$. Then applying (3.16) to the above we obtain, changing constants at will,

$$P(\sup_C|Lx| > \lambda) \leq cN^G(\theta_\lambda)\{(\sigma-\delta)^\beta\lambda^{-1}e^{-\lambda^2/2\sigma^2} + e^{-\lambda^2/2\delta^2}\}$$
$$+ c\, n_{\theta_\lambda}\lambda^{-1}e^{-\lambda^2/2\sigma^2} \ . \tag{5.13}$$

Choose $\delta = \sigma-\lambda^{-2}\ell n\lambda^{2\beta\sigma^3}$, taking λ large enough so that $\delta\in(\tfrac{1}{2}\sigma,\sigma)$, and note that for this δ

$$(\sigma-\delta)^\beta e^{-\lambda^2/2\sigma^2} \leq c\lambda^{-2\beta}(\ell n\lambda)^\beta e^{-\lambda^2/2\sigma^2} \ ,$$

and

$$e^{-\lambda^2/2\delta^2} = \exp\{-\frac{\lambda^2}{2\sigma^2}(1+\frac{\sigma^2-\delta^2}{\delta^2})\}$$

$$= \exp\{-\frac{\lambda^2}{2\sigma^2}[1+(\frac{\sigma+\delta}{\delta^2})\lambda^{-2}\ell n\lambda^{2\beta\sigma^3}]\}$$

$$\leq \exp\{-\frac{\lambda^2}{2\sigma^2} - \frac{2\beta\sigma}{\delta}\ell n\lambda\}$$

$$\leq \lambda^{-2\beta}e^{-\lambda^2/2\sigma^2} .$$

Substituting these last two inequalities into (5.13) establishes (3.17) and, thus, the theorem.

Proof of Theorem ?.5 We work from Corollary 3.3. For fixed λ define the sequence $\{\delta_i\}$ given by

$$\delta_0 = \tfrac{1}{2}\sigma, \quad \delta_i^2 = \sigma^2 - (m-i)\lambda^{-2}, \quad i=1,\ldots,m,$$

where $m := [\tfrac{1}{2}\sigma^2\lambda^2]$. Clearly it will suffice for us to bound $P(\sup_{C^+_{\delta_0}}|Lx| > \lambda)$. Apply Corollary 3.3 to obtain

$$P\{\sup_{C^+}|Lx| > \lambda\} \leq c \sum_{i=1}^{m} n(\delta_{i-1},\delta_i,\theta)\lambda^{-1}\exp(-\tfrac{1}{2}\lambda^2/\delta_i^2).$$

Note that $\delta_i - \delta_{i-1} \leq 1/(\sigma\lambda^2)$. Take λ large enough for (3.19) to hold, and substitute to bound the above sum by

$$\text{(5.14)} \qquad c[\sigma^{-\beta}\lambda^{-1-2\beta}N^{G}(\theta_\lambda) + n_{\theta_\lambda}\lambda^{-1}]\cdot \sum_{i=1}^{m}\exp(-\tfrac{1}{2}\lambda^2/\delta_i^2)$$

Thus to complete the proof we need only bound the last summation by $e^{-\frac{1}{2}\lambda^2/\sigma^2}$. This can be done as follows. Set

$$\alpha_i = \exp\{-\tfrac{1}{2}\lambda^2/(\sigma^2-(m-i)\lambda^{-2})\}$$

It is easy to check that

$$\alpha_{i-1} \le \alpha_i e^{-1/2\sigma^4} < \alpha_i \;.$$

Thus the sum in (5.14) is bounded by

$$\alpha_m \sum_{k=1}^{m} (e^{-1/2\sigma^4})^k \le \frac{\alpha_m}{1-e^{-1/2\sigma^4}} = c\, e^{-\frac{1}{2}\lambda^2/\sigma^2} \;,$$

which completes the proof.

Remark: The astute reader may have noticed that at no point in any of our proofs have we used the full power of the Basic Inequality (2.9), in that we have not taken advantage of the ε and λ double sequences to partition C according to variance (i.e. into the sets of G_θ). The reason is that, while doing so we can improve on the standard upper bounds, we cannot reach the sharpness of, say, Theorem 3.5 without an intermediate result like Theorem 3.3. In fact, it is the careful conditioning argument that goes into the proof of Theorem 3.3 that, ultimately, makes everything work.

6. SOME COMMENTS

1. Lower bounds. Throughout this paper we have, with the exception of Example 4.3 treating rectangle indexed sheets, dealt only with upper bounds for the excursion probability. The fact that in every example for which lower bounds are available we find that our upper bounds are sharp in the power of λ leads one to believe that they may be sharp in general. This, however, does not seem to be easy to prove. Some lower bounds are available from Weber (1980) and these, like his upper bounds, are sharp for processes with constant variance. For the highly nonstationary examples of Section 4 they do not provide bounds that match our upper bounds. As for upper bounds, however, it is easy to see by example that lower bounds that depend only on entropy without taking into consideration varying variance can never be sharp for all cases.

2. Vapnik-Cervonenkis classes. The natural geometric structure of VC classes of sets or functions should be enough to generate some of the homogeneity of C required by our theorems. Furthermore, the fact that each VC class has a natural, single parameter describing its structure (and, in a certain sense, its "dimensionality") seems to indicate that it should be possible to apply our results to VC classes in such a way that this parameter enters in a simple fashion into the power of λ. We have found indications that this should be true, but have been unable, so far, to put together a serious proof.

3. Exponential entropy. The exponent of entropy of C is defined by

$$T = T(C) = \limsup_{\varepsilon \downarrow 0} \log \log N(C,\varepsilon)/\log(1/\varepsilon).$$

For L to be continuous on C we must have $r \leq 2$. If $r < 2$, L is continuous. For $r=2$ there are examples of both continuous and discontinuous (and hence unbounded) processes. By assuming, as we have since Section 3, polynomial entropy, we assume $r=0$, thus leaving out many interesting examples. In particular, we cannot handle many set indexed sheet problems. (See Dudley (1973, 78) for examples.) Furthermore, bounds of the form $\lambda^{\alpha}e^{-\frac{1}{2}\lambda^2}$ are not valid in this case. Nevertheless, a result of Borell (1975, p. 214, middle of proof) states that for all a.s. bounded Gaussian processes there is a bound of the form $\exp(-\frac{1}{2}\lambda^2 + \text{const.}\,\lambda)$. In fact, Borell's result can be improved on, and, under mild conditions it is possible to show that there is a function $\alpha : [0,2] \to [0,1]$ for which bounds of the form $\exp(-\frac{1}{2}\lambda^2 + \text{const.}\,\lambda^{\alpha(r)})$ hold. We shall report this separately.

References

[1] Adler, R.J. (1984) The supremum of a particular Gaussian field. Ann. Probab. 12 436-444.

[2] Adler, R.J. and Brown, L.D. (1985) Tail behaviour for suprema of empirical processes. Ann. Probab. 13 to appear.

[3] Belyaev, Yu.K. and Piterbarg, V.I. (1972). The asymptotic formula for the mean number of A-points of excursions of Gaussian fields above high levels (in Russian). In Bursts of Random Fields, Moscow Univ. Press, Moscow. 62-89.

[4] Berman, S.M. (1974). Sojourns and extremes of Gaussian processes. Ann. Probab. 2 999-1026.

[5] Berman, S.M. (1985a). An asymptotic bound for the tail of the distribution of the maximum of a Gaussian process. Ann. Inst. Henri Poincaré 25 47-57.

[6] Berman, S.M. (1985b). As asymptotic formula for the distribution of the maximum of a Gaussian process with stationary increments. J. Appl. Prob. 22 454-460.

[7] Borell, C. (1975). The Brunn-Minkowski inequality in Gauss space. Inventiones Math., 30 207-216.

[8] Cabaña, E.M. (1984). On the transition density of a multidimensional parameter Wiener process with one barrier. J. Appl. Prob. 21 197-200.

[9] Cabaña, E.M. and Wschebor, M. (1981). An estimate for the tails of the distribution of the supremum for a class of stationary multi-parameter Gaussian processes. J. Appl. Prob. 18 536-541.

[10] Cabaña, E.M. and Wschebor, M. (1982). The two parameter Brownian bridge: Kolmogorov inequalities and upper and lower bounds for the distribution of the maximum. Ann. Prob. 10 289-302.

[11] Cressie, N. and Davis, R.W. (1981) The supremum distribution of another Gaussian process. J. Appl. Prob. 18 131-138.

[12] Darling, D.A. (1983). On the supremum of a certain Gaussian process. Ann. Probab. 11 803-806.

[13] Dudley, R.M. (1967). The sizes of compact sets of Hilbert space and continuity of Gaussian processes. J. Funct. Anal., 1 290-330.

[14] Dudley, R.M. (1973) Sample functions of the Gaussian process. Ann. Probab. 1 66-103.

[15] Dudley, R.M. (1978) Central limit theorems for empirical measures. Ann. Probab. 6 899-929.

[16] Dudley, R.M. (1984) A course on empirical processes. Ecole d'ete de Probabilities de St. Flour 1982. To appear in Lecture Notes in Math, Springer.

[17] Fernique, X. (1970). Intégrabilité des vecteurs gaussienns. C.R. Acad. Sci. Paris Ser A. 270 1698-1699.

[18] Fernique, X. (1975) Regularité des trajectoires des fonctions aleatoires gaussienns. in Ecole de Eté de Probabilitiés de Saint Flour IV, 1974. Lecture Notes in Math. 480, Springer.

[19] Goodman, V. (1976). Distribution estimates for functionals of the two-parameter Wiener process. Ann. Probab. 4 977-982.

[20] Goodman, V., Kuelbs, J. and Zinn, J. (1981) Some results on the LIL in Banach space with application to weighted empirical processes. Ann. Probab. 9 713-753.

[21] Kiefer, J. (1972). Skorohod embedding of multivariate R.V.'s and the sample D.F. Z. Wahrscheinlichkietstheorie verw. Gebiete. 24. 1-35.

[22] Kuelbs, J. (1975) Sample path behavior for Brownian motion in Banach space. Ann. Probab. 3 247-261.

[23] Landau, H.J. and Shepp, L.A. (1971). On the supremum of a Gaussian process. Sankyā Ser A32 369-378.

[24] Leadbetter, M.R., Lindgren, G. and Rootzen, H. (1983) Extremes and Related Properties of Random Sequences and Processes, Springer-Verlag, New York.

[25] Marcus, M.B. and Shepp, L.A. (1971) Sample behaviour of Gaussian processes. Proc. Sixth Berkeley Symp. Math. Statist. Probability II. Univ. of California Press, 423-441.

[26] Pickands, J. (1969a). Upcrossing probabilities for stationary Gaussian processes. Trans. Amer. Math. Soc. 145 51-73.

[27] Pickands, J. (1969b). Asymptotic properties of the maximum in a stationary Gaussian process. Trans. Amer. Math. Soc. 145 75-86.

[28] Piterbarg, V.I. and Prisjaznjuk, V.P. (1979). Asymptotics of the probability of large excursions for a nonstationary Gaussian process. Theory Probab. Math. Statist. 18 131-144.

[29] Segal, I.E. (1954) Distributions in Hilbert space and canonical systems of operators. Trans. Amer. Math. Soc. 88 12-41.

[30] Shepp, L. (1971). First passage time for a particular Gaussian process. Ann. Math. Statist. 42 946-951.

[31] Slepian, D. (1961) First passage time for a particular Gaussian process. Ann. Math. Statist. 32 610-612.

[32] Slepian, D. (1962). The one-sided barrier problem for Gaussian noise. Bell System Tech. J. 41 463-501.

[33] Slepian, D. and Shepp, L. (1976). First passage time for a particular stationary periodic Gaussian process. J. Appl. Prob. 13 27-38.

[34] Weber, M. (1978). Classes supérieures de processus gaussiens. Z. Wahrscheinlichkietstheorie verw. Gebiete 42 113-128.

[35] Weber, M. (1980). Analyse asymptotique des processus gaussiens stationnaires. Ann. Inst. Henri Poincaré 16 117-176.

END

DTIC

5-86